



Big Data and Dementia: Charting the Route Ahead for Research, Ethics, and Policy

Marcello Ienca, Effy Vayena and Alessandro Blasimme*

Health Ethics and Policy Laboratory, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

Emerging trends in pervasive computing and medical informatics are creating the possibility for large-scale collection, sharing, aggregation and analysis of unprecedented volumes of data, a phenomenon commonly known as big data. In this contribution, we review the existing scientific literature on big data approaches to dementia, as well as commercially available mobile-based applications in this domain. Our analysis suggests that big data approaches to dementia research and care hold promise for improving current preventive and predictive models, casting light on the etiology of the disease, enabling earlier diagnosis, optimizing resource allocation, and delivering more tailored treatments to patients with specific disease trajectories. Such promissory outlook, however, has not materialized yet, and raises a number of technical, scientific, ethical, and regulatory challenges. This paper provides an assessment of these challenges and charts the route ahead for research, ethics, and policy.

Keywords: dementia, big data, ethics, health policy, Alzheimer's disease, real-world evidence, mild cognitive impairment, m-health

OPEN ACCESS

Edited by:

Marco Canevelli,
Sapienza Università di Roma, Italy

Reviewed by:

Jagadish K. Chhetri,
Gérontopôle-Institute on Aging,
Toulouse University Hospital, France
Nagaendran Kandiah,
National Neuroscience Institute
(NNI), Singapore

*Correspondence:

Alessandro Blasimme
alessandro.blasimme@hest.ethz.ch

Specialty section:

This article was submitted
to Geriatric Medicine,
a section of the journal
Frontiers in Medicine

Received: 03 November 2017

Accepted: 16 January 2018

Published: 06 February 2018

Citation:

Ienca M, Vayena E and Blasimme A
(2018) Big Data and Dementia:
Charting the Route Ahead for
Research, Ethics, and Policy.
Front. Med. 5:13.
doi: 10.3389/fmed.2018.00013

INTRODUCTION: BIG DATA, HEALTH, AND DEMENTIA

The predicted threefold increase in the prevalence of dementia in the coming decades (1) will put health-care systems and informal caregivers under an unmanageable pressure. This explains the current eagerness to find a cure, to slow down progression, and to develop better predictive tools to spot early signs of cognitive decline that may develop into full-blown dementias. The identification of biomarkers for Alzheimer's disease (AD) has paved the way for research into early determinants, with an emphasis on treating patients before the clinical manifestation of the disease (2). According to a study of the drug development pipeline for AD in 2017, current efforts focus mainly on disease-modifying therapies (DMTs) (3), and most of DMT agents in phase III AD trials (as reported on <https://clinicaltrials.gov> in 2017) address amyloid targets (15 out of 18) (idem). Yet, the failure rate for these kinds of trials is notoriously high (4).

Increased tau protein in the cerebrospinal fluid (CSF) is also used as a biomarker for patients affected by mild cognitive impairment (MCI) who are likely to transition to AD (5). Still, the very use of MCI as a prodromal phase of AD has recently come under attack, as MCI itself reverts to normal cognitive functioning in many patients (6, 7).

Genome wide association studies have identified genes associated with increased risk of late onset AD (idem). In particular, APOe4 has been shown to have some, but only limited predictive value in determining the progression from (MCI) to AD (8).

Further efforts are underway to establish other biomarkers—genetic, biological, and combinations thereof—in both CSF and plasma (2), as well as attempts to look at cognitive decline on longitudinally longer scales (9).

Multidimensional models integrating multiple biomarker data and etiological pathways are regarded as decisively more promising than reductionist approaches based on single explanatory hypothesis like those driving clinical research on the amyloid cascade (4, 10).

Integrative approaches to dementias are ideal candidates to test an incipient approach driven by the use of big data in dementia research. Recently, the global surge of interest around big data has spilled over into the field of dementia, as well as in many other domains of medicine and health research. Big data refers to unprecedentedly large amounts of data analyzed through novel data mining techniques for a variety of different purposes. Although the concept lacks an agreed-upon definition, it is generally assumed that big data is characterized by the so-called 3Vs: volume, velocity, and variety (11). This simplified definition captures some peculiar facts. For starters, the total amount of data stored in data centers in the world is estimated to reach 915 exabytes by 2020 (2.4-fold increase with respect to today) (12). In parallel the amount of data generated by networked end-user devices and appliances will be more than triple in the same period (12). Data are being generated at impressive velocity. Both structured and unstructured data contribute to this rapid growth. Biomedicine does its part too. Sequencing a whole human genome now take as little as few hours, and it is estimated that more than 35 petabytes of genomic data are produced every year (13). But genomic data are not the only reason why health care is predicted to be one of the domains in which big data will have a transformative effect (14). Electronic health records (EHRs)—comprehensive records of patient health history in electronic format—have now entered routine clinical practice in most advanced countries. Sensor-equipped mobile devices, wearables, and appliances keep track of physical activity, location, sleeping habits, and vital parameters in real time, 24 h a day. Mobile-based applications (hereafter apps), loyalty cards, credit cards, and smart objects register accurate data about our consuming habits. Biomedical research produces huge amounts of data that scientists can store and share for secondary use through a variety of data repositories and biobanks. The ability of such heterogeneous information to provide a multidimensional account of one's health state has led some to speak of it as our "digital phenotype" (15), and to introduce the notion of "digital health" (16). Interest is rapidly growing around the potential medical breakthroughs enabled by mining such unprecedented amounts of data.

Dementia research is among those fields in which big data are regarded as more promising. Leveraging the collection, aggregation, and predictive analysis of large data volumes could reboot dementia research and care as it holds the potential of casting new light on its etiology, enabling timelier diagnosis and prevention strategies, and possibly overcoming current therapeutic limitations. In particular, this potential is more likely to be realized by enabling the integration of EHRs, molecular biomarkers, neuroimaging biomarkers, and mobile health (mHealth) data.

In this contribution, we review the existing scientific literature on big data approaches to dementia including both original research articles and commercially available mHealth applications. Based on this synthesis, we identify some major promises and challenges associated with big data trends in dementia

research and care and chart the route ahead for research, ethics, and policy (see **Figure 1**).

METHODS

In spite of their disruptive potential, big data trends in dementia still remain a relatively unexplored topic. In October 2017, we conducted a PubMed search [(“big data” AND “dementia”) OR (“big data” AND “cognitive decline”)] with unrestricted time range. The search retrieved nine peer-reviewed articles and one conference paper. The nine scientific articles included editorials (17, 18) and commentaries aimed at introducing big data principles to the dementia community, examining the applicability of these principles to dementia research or discussing the level of preparedness of the dementia community for big data approaches (19). While most articles focused primarily on AD dementia, a recent commentary explored the use of big data resources to optimize data use in vascular dementia research (20).

The data sources examined in these studies included EHRs (20), Internet searches (21), and genetic data (22). Additional data sources such as mHealth data did not appear in the foreground of current scientific literature. However, a parallel search on the two main digital distribution services (app stores), namely, Google Play and iTunes (the latter screened through API search), indicated the current availability of 35 mobile apps aimed at screening dementia and cognitive decline through self-assessment tools or digital assistants for health professionals.

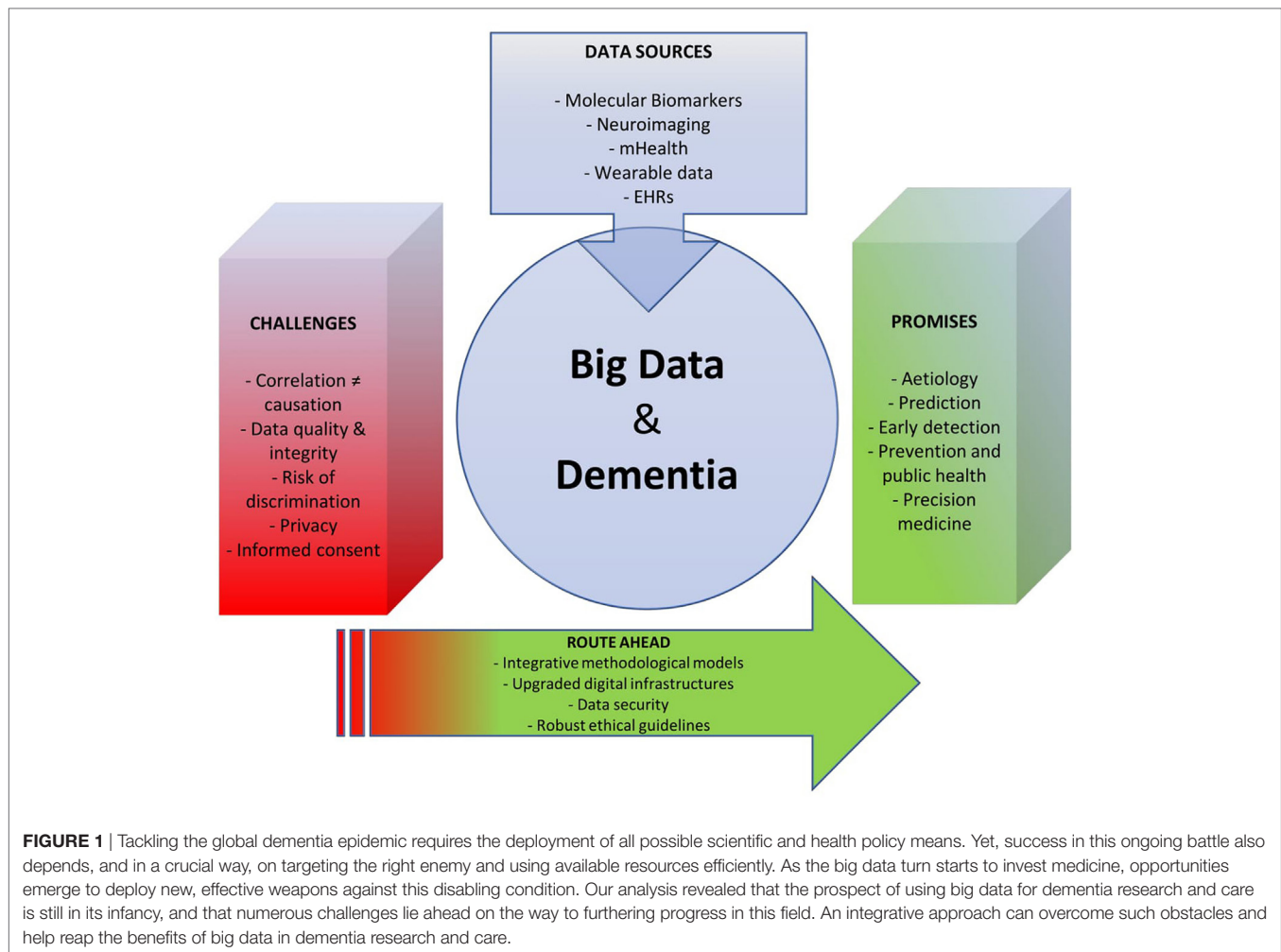
Further screening was performed through unstructured online search and retrieval of secondary sources. This second phase of analysis revealed that, besides traditional research models, big data approaches to dementia have also been pursued in the form of data analytics challenges. For instance, the DREAM Challenge for AD aimed at: (i) predicting changes in cognitive scores 2 years after initial assessment; (ii) understanding the biological basis of resilience to amyloid pathology; and (iii) classifying individuals into diagnostic groups using neuroimaging. The challenge capitalizes on large datasets such as the Alzheimer's Disease Neuroimaging Initiative's database and leverages multiple data sources including genetics, neuroimaging, cognitive assessment, and demographic information.

While research articles on big data and dementia still appear relatively rare, several policy documents produced by inter-governmental organizations focus on this topic. Between 2013 and 2015, the Organisation for Economic Co-operation and Development (OECD) released five policy papers addressing promises and challenges of big data trends in dementia care and research (23–27).

RESULTS

Promises

As often observed, dementia is “both a global problem and a pathological conundrum” (18). Therefore, the deployment of big data techniques should aim both at alleviating the global burden of dementia and at providing novel explanatory resources for the further understanding of the disease.



Big data may in the future help in cracking the pathological conundrum of AD by shedding new light on its etiology. The use of single molecular biomarkers in isolation has so far not successfully predicted the functional and cognitive outcomes of dementia (28). Based on this observation, researchers have criticized the disproportionate focus on single molecular biomarkers such as amyloid- β and tau in dementia research, and called for more integrative approaches to the study of the etiology (10, 29). For this reason, big data trends could corroborate a multi-modal (having different modes of data acquisition from heterogeneous data sources) and multi-scalar (from the molecular to the behavioral and population scale) account of dementia, enabling statistical associations across different data types and scales. This might fix the lack of integration that currently affects different data levels, hence helping to “glimpse the big picture” (30)—from genes and molecules to cognition and behavior—in dementia research.

In recent years, attempts to integrate and statistically correlate information across large population groups have been used to achieve earlier detection of AD, ideally in the pre-symptomatic phase. Researchers have utilized multiple data sources including

neuroimaging (31, 32), hand gesture tracking (33), and retinal scans (34), and developed machine learning algorithms to detect structural (e.g., changes in brain structure) or functional (e.g., changes in time response during the completion of tasks) anomalies years prior to the onset of symptoms. Some of these studies show prognostic and predictive potential (32, 33). In parallel, a number of smartphone-based apps claim the ability to detect early signs of cognitive decline through gaming behavior or digitalized mental examination. For example, the mobile app *Sea Hero Quest* (35) gamified virtual spatial navigation assessment techniques (36) to find correlations between levels of gaming performance and cognitive decline (37). With over 2.5 million users, the app can capitalize on larger datasets than most conventional clinical tools, hence leveraging the potential of big data analytics to improve screening and early detection. While *Sea Hero Quest* is designed for self-assessment, other apps such as *CogniSense* (38) and *ACEmobile* (39) provide digital assistance to clinicians during the cognitive screening of patients.

Employing big data to reduce the global burden of dementia (25) implies aggregating large sets of population-scale data to improve prevention and public health strategies. For example,

Doubal et al. have argued that the linkage of routine clinical data at a national or even international level might result in a better understanding of the risk factors of dementia, resulting in more effective prevention (20). In a similar fashion, Dacks et al. have argued that optimizing the use of observational data through data-pooling and Internet-based tools might support personal, clinical, or public health decision-making and contribute to the development of specific interventions that reduce modifiable risk factors (40). In parallel, as stated in an OECD report, integrating data from different size scales (from molecular to whole population) and information types (from neurophysiological to behavioral) might enable the development of precision medicine solutions (24). With recent unmet expectations regarding pharmacological therapy, this *personalized turn* can increase therapeutic effectiveness through the customization of treatments and medical decisions to each individual patient (41, 42). This is particularly relevant as evidence suggests that the “biological processes driving the clinical phenotype can differ remarkably from patient to patient” (28) and can also depend on comorbidities, co-medications, and patient genotypes.

Challenges

Big data and digital health tools can streamline the detection of early signs of cognitive decline. MCI is a prominent example of a prodromal syndrome that, while not strongly debilitating, can develop into full-blown dementia (43).

Other entities have recently been suggested to have a similar predictive value, such as subjective cognitive decline and mild behavioral impairment (44, 45). Apps that monitor the clinical presentation of such conditions (such as users' mood and cognitive performance over time) can greatly facilitate spotting changes from baseline functioning levels. The ubiquity and ease-of-use of such tools may enable self-monitoring practices, thus leading to a widespread increase in the number of patients referred to neurologists. On the one hand, if this enhances early detection of patients at risk of developing dementia, such practices can inflate the number of false positive diagnoses. This is a concrete risk, especially in light of recent findings showing that conditions like MCI, for instance, are prone to stabilization or even spontaneous reversion to normal functioning (46). In the absence of adequate evidence and guidelines, excessive emphasis on preclinical syndromes can lead to over-diagnosis and unnecessary medicalization, with consequences for both health systems and patients. The latter face risks of insurance and employment discrimination, stigmatization, and direct psychological harm. Health systems, on the other hand, may have to cope with unnecessary financial and organizational costs in response to an upsurge of mild/preclinical syndromes (46–48).

These downstream effects of mHealth applications in dementia demonstrate the need to establish *ad hoc* regulatory pathways to validate apps that make medical claims. Regulators such as the US Food and Drugs Administration (49) and the UK Medicines and Healthcare Products Regulatory Agency (50) are starting to venture in this domain. Nonetheless, we observed that many developers present their applications in ambiguous ways, without offering sufficient information regarding either the evidence backing their products, or the way in which they will handle the

personal data of their customers. More effective incentives and disincentives are still needed to ensure sufficient consumer protection in this area. In particular, there needs to be more stringent oversight regarding the health-related claims of apps.

Despite these drawbacks, early detection remains laudable, as it allows us to treat patients before degenerative trajectories compromise the odds of slowing or stabilizing cognitive decline. The collection and analysis of multiple data types describing the aging trajectory of individual patients can help isolate discrete stressors and molecular characteristics (including genetic ones), and cluster them with cognitive or neuro-psychiatric symptoms that lead to the clinical manifestation of dementia. This kind of knowledge will improve the clinical understanding of dementia, as well as the development of personalized therapeutic and preventive interventions. Creating the evidence base for such interventions will require large-scale personal data repositories, giving rise to regulatory challenges in terms of data management, protection, aggregation, interoperability, privacy, and informed consent for the collection, use, and sharing of such data. Such issues are currently being debated in ethics and regulatory circles (51). Yet, other pressing issues relate to the most adequate means to generate clinically reliable and usable evidence from heterogeneous “real-world data,” such as EHRs, mobile device data, and socioeconomic data (52). Current discussions on pragmatic trial designs will likely turn out relevant to research in novel, data-driven approaches to prevention and therapy in dementia (53). In the case of a stigmatizing condition like dementia linking data from within and without the clinical setting to detect early signs of cognitive decline poses peculiar ethical challenges. At a minimum, patients (or their legal representatives, in case of advanced dementia) should be made aware of such activities and be given the option to opt out.

A WAY FORWARD: THE NEED FOR AN INTEGRATIVE APPROACH

Scientific Evidence and Theory

To overcome the above structural challenges and ensure the success of big data initiatives for dementia, there is a need for integrative approaches at the level of research methodology, digital infrastructures, and financing, as well as ethics and regulation.

At the scientific level, there is a need for clearly demarcating the explanatory power of big data driven research. Large-scale data collection and further mining through analytical tools could boost dementia research and care management, establishing reliable statistical correlations between heterogeneous data sources whose association could not be detected through small-scale methods. It is questionable, however, whether big data approaches alone would suffice to uncover the causal mechanisms of AD pathology. The idea that large datasets might *speak for themselves* (54), independent of explanatory hypotheses (55) has attracted praise as well as criticism (56, 57). Integrative approaches are needed to combine the predictive power of big data with theoretically robust and causally explanatory scientific models. A valid proposal for such integration has been advanced by Geerts

et al., who suggested that data-driven analytic approaches in AD research “need to be organically integrated into a quantitative understanding of the pathology” involving mechanism-based modeling and simulation approaches. In their view, this integration could enable a shift from big data to *smart-data*, i.e., from “information” to “actionable knowledge” (28). Similarly, DeKosky has called for the integration of big data approaches with basic neuroscience (18).

Integrative and theory-mediated big data approaches are well placed to overcome current limitations in dementia research. Taking a stance from systems biology and complexity theory, Geerts et al. argued that the integration of big data analytics with modeling and simulation might overcome the explanatory failures of reductionist biological approaches focused on single molecular biomarkers in isolation. In their view, this holistic approach has already shown benefits in PD-dementia, as it has led to a better understanding and optimization of deep-brain stimulation protocols (28).

Digital Infrastructure

Current digital infrastructures of dementia research and care need to be upgraded. As stated in a recent OECD report, secure infrastructures for data storage, processing, and access need to be sustained through complementary resources (27). While ongoing digitalization and automation in dementia care offer novel opportunities for large-scale data acquisition (58), further efforts are required to sustain the secure and reliable sharing of such information, and to guarantee the interoperability of different data-repositories (e.g., genetic biobanks, neuroimaging repositories, and eHealth data platforms). Active cooperation between public and private actors has been recognized as a viable strategy for increasing funding opportunities and favoring the digital transformation of dementia research and care (26). Yet, the appropriation of health data by large ICT corporation can cause public unease, thus undermining the development of data-driven medicine. A recent controversy over Google’s access to NHS data through its AI subsidiary DeepMind shows that sufficient safeguards are not yet in place (59).

Ethical Guidelines

Ethically robust guidelines for the collection and sharing of personal health data would facilitate big data research while maintaining public trust and protecting data subjects. Existing oversight mechanisms (such as ethics review) and conventional informed consent models appear “ill suited” for large-scale data collections (60, 61). As far as research on large-scale repositories of structured and unstructured data is concerned, *ad hoc* criteria for assessing research protocols employing novel data analytics

tools are urgently needed. Moreover, informed consent—in its current shape—does not grant data subjects (nor their legal representatives) sufficient control over highly sensible information regarding their cognitive state. This is a disincentive for people to make their data available for research in the first place. New mechanisms are being explored to enable more granular data control on the part of data subjects—for example by implementing data portability rights, or by introducing electronic consent management mechanisms and participatory forms of data governance (61). The longitudinal nature of studies based on big data and real-world evidence, however, poses issues relative to the validity of initial consent obtained from people whose cognitive functions may degrade over time. *Ad hoc* oversight mechanisms, such as advanced directives, shall be in place to safeguard the autonomy and wellbeing of those patients while at the same time enabling scientific progress.

Privacy and Data Security

As to privacy and data security, researchers and regulators need to acknowledge that not even anonymization of deep phenotypic data—such as those that are needed for the development of big data approaches in dementia—is a sufficient firewall. As demonstrated in the case of genomic (62) and electrophysiological data (63), maliciously re-identifying anonymous data is within reach of sufficiently skilled offenders. Given the sensitive nature of data regarding cognition (64) and given the frequency of spectacular health data breaches (65), data security represents a priority. It follows that efficient regulatory and technical measures to shore up data security are key to scientific progress in the field. Privacy-preserving techniques such as encryption and block chain need to be incorporated into the digital infrastructure of current data transmissions. These solutions will not only enhance data security but also facilitate and sustain the trust of individuals (both healthy subjects and people with symptomatic dementia) in the data collection and sharing (66).

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by the Swiss National Science Foundation under awards 407540_167223 (NRP 75 Big Data) and PP00P3_157556, and the Swiss Academy of Medical Sciences under award KZS 20/17.

REFERENCES

1. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement* (2013) 9:63–75.e2. doi:10.1016/j.jalz.2012.11.007
2. Huynh RA, Mohan C. Alzheimer’s disease: biomarkers in the genome, blood, and cerebrospinal fluid. *Front Neurol* (2017) 8:102. doi:10.3389/fneur.2017.00102
3. Cummings J, Lee G, Mortsdorf T, Ritter A, Zhong K. Alzheimer’s disease drug development pipeline: 2017. *Alzheimers Dement* (2017) 3:367–84. doi:10.1016/j.trci.2017.05.002

4. Canevelli M, Bruno G, Cesari M. The sterile controversy on the amyloid cascade hypothesis. *Neurosci Biobehav Rev* (2017) 83:472–3. doi:10.1016/j.neubiorev.2017.09.015
5. Vos SJ, Van Rossum IA, Verhey F, Knol DL, Soininen H, Wahlund L-O, et al. Prediction of Alzheimer disease in subjects with amnesic and nonamnesic MCI. *Neurology* (2013) 80:1124–32. doi:10.1212/WNL.0b013e318288690c
6. Canevelli M, Grande G, Lacorte E, Quarchioni E, Cesari M, Mariani C, et al. Spontaneous reversion of mild cognitive impairment to normal cognition: a systematic review of literature and meta-analysis. *J Am Med Dir Assoc* (2016) 17:943–8. doi:10.1016/j.jamda.2016.06.020

7. Canevelli M, Blasimme A, Vanacore N, Bruno G, Cesari M. From evidence to action: promoting a multidimensional approach to mild cognitive impairment. *J Am Med Dir Assoc* (2015) 16:710–1. doi:10.1016/j.jamda.2015.04.013
8. Elias-Sonnenschein LS, Viechtbauer W, Ramakers IH, Verhey FR, Visser PJ. Predictive value of APOE-epsilon4 allele for progression from MCI to AD-type dementia: a meta-analysis. *J Neurol Neurosurg Psychiatry* (2011) 82:1149–56. doi:10.1136/jnnp.2010.231555
9. Vos SJ, Xiong C, Visser PJ, Jasielec MS, Hassenstab J, Grant EA, et al. Pre-clinical Alzheimer's disease and its outcome: a longitudinal cohort study. *Lancet Neurol* (2013) 12:957–65. doi:10.1016/S1474-4422(13)70194-7
10. Herrup K. The case for rejecting the amyloid cascade hypothesis. *Nat Neurosci* (2015) 18:794–9. doi:10.1038/nn.4017
11. Laney D. 3D data management: controlling data volume, velocity and variety. *META Group Res Note* (2001) 6:70.
12. Cisco. *Cisco Global Cloud Index: Forecast and Methodology, 2015-2020, White Paper*. Cisco Public (2016).
13. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. Big data: astronomical or genetical? *PLoS Biol* (2015) 13:e1002195. doi:10.1371/journal.pbio.1002195
14. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute (2011).
15. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol* (2015) 33:462–3. doi:10.1038/nbt.3223
16. FDA. *Digital Health, U.S. Food and Drug Administration*. Silver Spring, MD: U.S. Department of Health & Human Services, Food & Drug Administration (2016).
17. Khachaturian AS, Meranus DH, Kukull WA, Khachaturian ZS. Big data, aging, and dementia: pathways for international harmonization on data sharing. *Alzheimers Dement* (2013) 9:S61–2. doi:10.1016/j.jalz.2013.09.001
18. DeKosky ST. The role of big data in understanding late-life cognitive decline: E Unum, Pluribus. *JAMA Neurol* (2014) 71:1476–8. doi:10.1001/jamaneurol.2014.2818
19. Hofmann-Apitius M. Is dementia research ready for big data approaches? *BMC Med* (2015) 13:145. doi:10.1186/s12916-015-0367-7
20. Doubal FN, Ali M, Batty GD, Charidimou A, Eriksdotter M, Hofmann-Apitius M, et al. Big data and data repurposing—using existing data to answer new questions in vascular dementia research. *BMC Neurol* (2017) 17:72. doi:10.1186/s12883-017-0841-2
21. Wang H-W, Chen D-R, Yu H-W, Chen Y-M. Forecasting the incidence of dementia and dementia-related outpatient visits with Google trends: evidence from Taiwan. *J Med Internet Res* (2015) 17:e264. doi:10.2196/jmir.4516
22. Swarup V, Geschwind DH. Alzheimer's disease: from big data to mechanism. *Nature* (2013) 500:34–5. doi:10.1038/nature12457
23. Deetjen U, Meyer ET, Schroeder R. *Big Data for Advancing Dementia Research: An Evaluation of Data Sharing Practices in Research on Age-Related Neurodegenerative Diseases*. Paris: OECD Publishing (2015).
24. Geoff A, Jillian O. *Dementia Research and Care Can Big Data Help? Can Big Data Help?* Paris: OECD Publishing (2015).
25. OECD. *Unleashing the Power of Big Data for Alzheimer's Disease and Dementia Research: Main Points of the OECD Expert Consultation on Unlocking Global Collaboration to Accelerate Innovation for Alzheimer's Disease and Dementia*. Paris: OECD Publishing (2014).
26. OECD. *Public-Private Partnerships in Biomedical Research and Health Innovation for Alzheimer's Disease and Other Dementias*. Paris: OECD Publishing (2015).
27. OECD. *Addressing Dementia*. Paris: OECD Publishing (2015).
28. Geerts H, Dacks PA, Devanarayan V, Haas M, Khachaturian ZS, Gordon ME, et al. Big data to smart data in Alzheimer's disease: the brain health modeling initiative to foster actionable knowledge. *Alzheimers Dement* (2016) 12:1014–21. doi:10.1016/j.jalz.2016.04.008
29. Ridler C. Dementia prevention hits the headlines at AAIC17. *Nat Rev Neurol* (2017) 13:576–576. doi:10.1038/nrneuro.2017.115
30. Husain M. Big data: could it ever cure Alzheimer's disease? *Brain* (2014) 137:2623–4. doi:10.1093/brain/awu245
31. Amoroso N, La Rocca M, Bruno S, Maggipinto T, Monaco A, Bellotti R, et al. *Brain Structural Connectivity Atrophy in Alzheimer's Disease*. arXiv preprint arXiv:1709.02369. (2017).
32. Mathotaarachchi S, Pascoal TA, Shin M, Benedet AL, Kang MS, Beaudry T, et al. Identifying incipient dementia individuals using machine learning and amyloid imaging. *Neurobiol Aging* (2017) 59:80–90. doi:10.1016/j.neurobiolaging.2017.06.027
33. Souillard-Mandar W, Davis R, Rudin C, Au R, Libon DJ, Swenson R, et al. Learning classification models of cognitive conditions from subtle behaviors in the digital clock drawing test. *Mach Learn* (2016) 102:393–441. doi:10.1007/s10994-015-5529-5
34. Koronyo Y, Biggs D, Barron E, Boyer DS, Pearlman JA, Au WJ, et al. Retinal amyloid pathology and proof-of-concept imaging trial in Alzheimer's disease. *JCI Insight* (2017) 2:1–19. doi:10.1172/jci.insight.93621
35. Morgan J. Gaming for dementia research: a quest to save the brain. *Lancet Neurol* (2016) 15:1313. doi:10.1016/S1474-4422(16)30123-5
36. Ventura M, Shute V, Wright T, Zhao W. An investigation of the validity of the virtual spatial navigation assessment. *Front Psychol* (2013) 4:852. doi:10.3389/fpsyg.2013.00852
37. Gholipour B. *Can You Diagnose Dementia from a Gaming App?* Scientific American (2016).
38. Clionsky M, Clionsky E. Dementia and the brain-breathing connection. *J Alzheimers Dis Parkinsonism* (2016) 6:e135. doi:10.4172/2161-0460.1000e135
39. Newman C, Hodges J, Pearson S, Noad R. The design and implementation of a computer supported assessment of dementia-ACEmobile. *Int J Integr Care* (2014) 14. doi:10.5334/ijic.1784
40. Dacks PA, Andrieu S, Blacker D, Carman AJ, Green AM, Grodstein F, et al. Dementia prevention: optimizing the use of observational data for personal, clinical, and public health decision-making. *J Prev Alzheimers Dis* (2014) 1:117–23.
41. Blasimme A, Vayena E. Becoming partners, retaining autonomy: ethical considerations on the development of precision medicine. *BMC Med Ethics* (2016) 17:67. doi:10.1186/s12910-016-0149-6
42. Blasimme A, Vayena E. “Tailored-to-you”: public engagement and the political legitimization of precision medicine. *Perspect Biol Med* (2017) 59:172–88. doi:10.1353/pbm.2017.0002
43. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA* (2014) 312:2551–61. doi:10.1001/jama.2014.13806
44. Jessen F, Amariglio RE, van Boxtel M, Breteler M, Ceccaldi M, Chételat G, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement* (2014) 10:844–52. doi:10.1016/j.jalz.2014.01.001
45. Ismail Z, Smith EE, Geda Y, Sultzer D, Brodaty H, Smith G, et al. Neuropsychiatric symptoms as early manifestations of emergent dementia: provisional diagnostic criteria for mild behavioral impairment. *Alzheimers Dement* (2016) 12:195–202. doi:10.1016/j.jalz.2015.05.017
46. Canevelli M, Bruno G, Remiddi F, Vico C, Lacorte E, Vanacore N, et al. Spontaneous reversion of clinical conditions measuring the risk profile of the individual: from frailty to mild cognitive impairment. *Front Med* (2017) 4:184. doi:10.3389/fmed.2017.00184
47. Canevelli M, Blasimme A, Vanacore N, Bruno G, Cesari M. Mild behavioral impairment: ethical, methodological and clinical reflections. *Neurosci Biobehav Rev* (2016) 69:402–3. doi:10.1016/j.neubiorev.2016.08.025
48. Canevelli M, Blasimme A, Vanacore N, Bruno G, Cesari M. Issues about the use of subjective cognitive decline in Alzheimer's disease research. *Alzheimers Dement* (2014) 10:881–2. doi:10.1016/j.jalz.2014.07.154
49. Hamel MB, Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. *N Engl J Med* (2014) 371:372. doi:10.1056/NEJMHle1403384
50. MHRA. *Medical Device Stand-Alone Software Including Apps*. London, UK: Drug Safety, Medicines and Healthcare products Regulatory Agency (2014). p. 1–38.
51. Vayena E, Gasser U, Wood A, O'Brien DR, Altman M. Elements of a new ethical framework for big data research. *Wash Lee L Rev Online* (2015) 72:423.
52. Jarow JP, LaVange L, Woodcock J. Multidimensional evidence generation and FDA regulatory decision making: defining and using “real-world” data. *JAMA* (2017) 318:703–4. doi:10.1001/jama.2017.9991
53. Kalkman S, van Thiel GJM, Grobbee DE, Meinecke A-K, Zuidgeest MGP, van Delden JJM. Stakeholders' views on the ethical challenges of pragmatic trials investigating pharmaceutical drugs. *Trials* (2016) 17:419. doi:10.1186/s13063-016-1546-3

54. Cukier K, Mayer-Schönberger V. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray (2013).
55. Anderson C. The end of theory: the data deluge makes the scientific method obsolete. *Wired Mag* (2008) 16:16–7.
56. Mazzocchi F. Could big data be the end of theory in science? A few remarks on the epistemology of data-driven science. *EMBO Rep* (2015) 16:1250–5. doi:10.15252/embr.201541001
57. Kitchin R. Big data, new epistemologies and paradigm shifts. *Big Data Soc* (2014) 1:2053951714528481. doi:10.1177/2053951714528481
58. Ienca M, Fabrice J, Elger B, Caon M, Pappagallo AS, Kressig RW, et al. Intelligent assistive technology for Alzheimer's disease and other dementias: a systematic review. *J Alzheimers Dis* (2017) 56:1301–40. doi:10.3233/JAD-161037
59. Hawkes N. NHS data sharing deal with Google prompts concern. *BMJ* (2016) 353:i2573. doi:10.1136/bmj.i2573
60. Vayena E, Mastroianni A, Kahn J. Caught in the web: informed consent for online health research. *Sci Transl Med* (2013) 5:173fs6. doi:10.1126/scitranslmed.3004798
61. Vayena E, Blasimme A. Biomedical big data: new models of control over access, use and governance. *J Bioeth Inq* (2017) 14:501–13. doi:10.1007/s11673-017-9809-6
62. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. *Science* (2013) 339:321. doi:10.1126/science.1229566
63. Ienca M, Haselager P. Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity. *Ethics Inf Technol* (2016) 18:117–29. doi:10.1007/s10676-016-9398-9
64. Ienca M, Andorno R. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sci Soc Policy* (2017) 13:5. doi:10.1186/s40504-017-0050-1
65. Perakslis ED, Stanley M. A cybersecurity primer for translational research. *Sci Transl Med* (2016) 8:322s2. doi:10.1126/scitranslmed.aaa4493
66. Adjekum A, Ienca M, Vayena E. What is trust? Ethics and risk governance in precision medicine and predictive analytics. *OMICS* (2017) 21:704–10. doi:10.1089/omi.2017.0156

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Ienca, Vayena and Blasimme. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.